

BEST AVAILABLE COPY

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 3 年 1 2 月 2 4 日
Date of Application:

出 願 番 号 特 願 2 0 0 3 - 4 2 6 1 6 8
Application Number:
[ST. 10/C]: [J P 2 0 0 3 - 4 2 6 1 6 8]

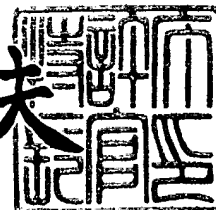
出 願 人 株式会社日立製作所
Applicant(s):

CERTIFIED COPY OF
PRIORITY DOCUMENT

2 0 0 4 年 2 月 1 6 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫



出証番号 出証特 2 0 0 4 - 3 0 0 9 4 6 4

【書類名】 特許願
【整理番号】 K03013061A
【あて先】 特許庁長官殿
【国際特許分類】 G06F 13/00
【発明者】
 【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所
 ソフトウェア事業部内
 【氏名】 亀田 正美
【特許出願人】
 【識別番号】 000005108
 【氏名又は名称】 株式会社 日立製作所
【代理人】
 【識別番号】 100075096
 【弁理士】
 【氏名又は名称】 作田 康夫
【選任した代理人】
 【識別番号】 100100310
 【弁理士】
 【氏名又は名称】 井上 学
【手数料の表示】
 【予納台帳番号】 013088
 【納付金額】 21,000円
【提出物件の目録】
 【物件名】 特許請求の範囲 1
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1

【書類名】 特許請求の範囲**【請求項 1】**

第1の記憶部から第2の記憶部へネットワークを介してデータ転送を行う計算機システムであって、前記第1の記憶部に格納されたデータをブロック転送プロトコルを用いて前記第2の記憶部へ転送する処理を行う第1の制御部と、複数のブロックのデータからなるファイルと当該ファイルを構成するデータのブロックとを対応付けて管理するテーブルと、前記第1の制御部からブロックを特定する情報を受け、前記テーブルを用いて当該ブロックに対応するファイルを特定し、当該特定されたファイルをファイル転送プロトコルを用いて前記第2の記憶部へ転送する処理を行う第2の制御部とを備えたことを特徴とする計算機システム。

【請求項 2】

前記ブロック転送プロトコルを用いた転送はS A N、前記ファイル転送プロトコルを用いた転送はL A Nを介して行うことを特徴とする請求項 1 記載の計算機システム。

【請求項 3】

前記第1の制御部は、前記第1の記憶部に格納されたデータをブロック転送プロトコルを用いて転送する際に、転送の障害を検知した場合、当該転送障害に関するブロックを特定する情報を前記第2の制御部へ通知することを特徴とする請求項 1 記載の計算機システム。

【請求項 4】

前記特定されたファイルは、前記転送障害に関するブロック以外のブロックのデータを含むことを特徴とする請求項 3 記載の計算機システム。

【請求項 5】

前記転送障害に関するブロック以外のブロックのデータとは、前記第1の制御部によるブロック転送プロトコルを用いた転送処理が完了したデータであることを特徴とする請求項 4 記載の計算機システム。

【請求項 6】

第1の記憶部から第2の記憶部へ第1、第2のネットワークを介してデータ転送を行う計算機システムであって、前記第1の記憶部に格納されたデータを第1のネットワークを介してブロック単位で前記第2の記憶部へ転送する第1の制御部と、前記第1の記憶部に格納されたデータを第2のネットワークを介してファイル単位で前記第2の記憶部へ転送する第2の制御部とを備え、前記第2の制御部は、複数のブロックのデータからなるファイルと当該ファイルを構成するデータのブロックとの関連付けを管理し、前記第1の制御部からのブロックを特定する情報を受け、当該ブロックのデータを含むファイルをファイル単位で前記第2の記憶部へ転送することを特徴とする計算機システム。

【請求項 7】

前記第2の制御部は、前記ファイル単位での転送処理に失敗した場合、当該転送処理に失敗したファイルに関連する複数のブロックを特定し、当該複数のブロックのデータ転送を前記第1の制御部へ指示することを特徴とする請求項 6 記載の計算機システム。

【請求項 8】

前記第1の記憶部は、同じ内容のデータが格納された正ボリュームと副ボリュームとからなり、前記第1の制御部は、前記副ボリュームに格納されたデータのブロック単位での転送処理に失敗した場合、当該転送処理に失敗したデータのブロックを特定する情報を前記第2の制御部へ通知し、前記第2の制御部から前記転送処理に失敗したファイルに関連する複数のブロックのデータ転送指示を受けて、前記正ボリュームに格納された前記複数のブロックに対応するデータのブロック単位での転送処理を行うことを特徴とする請求項 7 記載の計算機システム。

【請求項 9】

前記第1の制御部と前記第2の制御部とは同一の筐体にあることを特徴とする請求項 6 記載の計算機システム。

【請求項 1 0】

記憶部と当該記憶部に格納されたデータをブロックアドレスを用いてブロック単位に管理する第1の制御部とからなるストレージシステムに接続され、複数の前記ブロックアドレスを特定する情報とファイル識別子とを対応付け、複数のブロックからなるファイルをファイル単位に管理する第2の制御部を有する計算機システムにおける他の計算機システムへのデータ転送方法であって、前記第2の制御部は、前記第1の制御部から前記ブロックアドレスを特定する情報を受け、前記ブロックアドレスを特定する情報と対応付けられたファイル識別子を特定し、当該ファイル識別子に対応付けられた複数のブロックアドレスを特定する情報を前記第1の制御部へ通知し、当該第1の制御部から前記複数のブロックアドレスを特定する情報に対応するデータを受け、当該データに前記ファイル識別子を付してファイル単位で前記他の計算機システムへデータ転送することを特徴とするデータ転送方法。

【請求項11】

前記第2の制御部は、前記ブロックアドレスを特定する情報とファイル識別子とが対応付けられた管理テーブルを、ファイル単位でのデータ転送の際に前記他の計算機へ転送することを特徴とする請求項10記載のデータ転送方法。

【請求項12】

前記ブロックアドレスを特定する情報は、論理的なブロックアドレスであることを特徴とする請求項10記載のデータ転送方法。

【請求項13】

前記第1の制御部は、前記他の計算機システムに接続されたストレージシステムへのブロック単位でのデータ転送の障害を検知した場合、前記第2の制御部へ前記ブロックアドレスを特定する情報を通知することを特徴とする請求項10記載のデータ転送方法。

【請求項14】

前記計算機システムは、前記第1の制御部へブロックアドレスを特定する情報を通知してブロック単位でのデータ転送を依頼することを特徴とする請求項10記載のデータ転送方法。

【請求項15】

前記ブロック単位でのデータ転送とファイル単位でのデータ転送とは異なるネットワークを介して行われることを特徴とする請求項10記載のデータ転送方法。

【請求項16】

記憶領域と当該記憶領域に格納されたデータをブロックアドレスを用いてブロック単位に管理するコントローラとからなるシステムにファイバチャネルを介して接続され、複数の前記ブロックアドレスを特定する情報とファイル識別子とを対応付け、複数のブロックからなるファイルをファイル単位に管理するファイルサーバに、他のファイルサーバへのデータ転送方法を実行させるためのプログラムであって、前記コントローラから前記ブロックアドレスを特定する情報を受け、前記ブロックアドレスを特定する情報と対応付けられたファイル識別子を特定し、当該ファイル識別子に対応付けられた複数のブロックアドレスを特定する情報を前記コントローラへ通知し、当該コントローラから前記複数のブロックアドレスを特定する情報に対応するデータを受け、当該データに前記ファイル識別子を付してファイル単位で前記他のファイルサーバへデータ転送することを実行させるためのプログラム。

【書類名】 明細書**【発明の名称】** データ転送方法及びシステム並びにプログラム**【技術分野】****【0001】**

本発明は、ストレージシステム等の計算機システムに格納されたデータの転送技術に係り、特にNASその他のファイルシステムが有する機能を利用して、ストレージシステムに格納されたデータをバックアップ系のシステムに転送する技術に関する。

【背景技術】**【0002】**

現在、インターネット等のネットワーク環境の進展、デジタル情報の多様化（文書・図面・画像・ビジュアルコンテンツ等）により、情報処理システムが扱うデータの量が増加している。こうした大量のデータを取り扱うための情報処理システムは、大別すると、サーバ機能とストレージ機能とを有する。

【0003】

サーバ機能は、当該サーバ機能を有するシステム上で業務アプリケーション等を実行することにより、情報処理システムが提供する業務の処理を実行する。ストレージ機能は、サーバ機能を有するシステムが利用するデータを格納する。ストレージ機能を有するシステム、例えばストレージシステムには、サーバ機能を有するシステム、例えばファイルサーバごとに、ファイルサーバが認識できるファイルシステムでフォーマットされたボリュームが設けられている。ここで、ファイルサーバとストレージシステムは、SAN（Storage Area Network）と呼ばれるネットワークで接続することができる。SANは、通常、ファイバチャネル（Fibre Channel）と呼ばれる伝送路を用いて、一般的なネットワークと比較すると、高速・高信頼なデータ転送方式を実現している。SANは、複数のサーバとストレージシステムとを接続することができ、サーバで実行されている業務処理に影響を与えることなく、ストレージシステム間でデータのコピーを実行することも可能である。

【0004】

ここで、ストレージシステムに格納されているデータを、事故・災害・故障等から保護するため、別のストレージシステムや別ボリュームに転送及びコピーすることを「バックアップ」と呼ぶ。バックアップのためのサーバやストレージシステムは、WAN（Wide Area Network）を介して遠隔地に置かれることがある。

【0005】

上述の業務を処理するサーバ及びそのデータを格納するストレージシステムを、それぞれ現用系のサーバ、現用系のストレージシステムと呼び、バックアップのためのサーバ及びストレージシステムを、それぞれバックアップ系のサーバ、バックアップ系のストレージシステムと呼ぶ。バックアップ系のサーバとストレージシステムは、現用系と同様に、SANで接続されている。

【0006】

サーバは、上述のSANとは別のネットワークであるLAN（Local Area Network）にも接続されている。LANでは、IP（Internet Protocol）と呼ばれる通信方式で、データの送受信が行われる。

【0007】

従来の技術では、バックアップ時は、現用系のストレージシステムのデータを、SANとWANを介して、バックアップ系のストレージシステムにディスクのブロック単位でデータを転送する（例えば、特許文献1参照）。データ転送中、SANとWANを介した経路上で何らの障害が発生すると、データ転送は中止されることになる。

【0008】

また、LAN及びWANを経由して、IPを用いて現用系のデータをバックアップ系にデータ転送することも考えられる。一般的なIPを使用したデータ転送では、転送経路の候補となる複数のネットワークが存在する状況で、あるネットワークに障害がある場合

に、他のネットワークに切り替えるという従来技術がある（例えば、特許文献2参照）。

【0009】

さらに、通信プロトコルが異なる複数のネットワークから転送経路を選択する従来技術がある（例えば、特許文献3参照）。

【0010】

【特許文献1】特開2002-7304号公報

【0011】

【特許文献2】特開2001-197112号公報

【特許文献3】特開2003-32290号公報

【発明の開示】

【発明が解決しようとする課題】

【0012】

SANを介したバックアップを行う従来技術（特許文献1）では、現用系のストレージシステムのデータを、SAN及びWANを経由してバックアップ系に転送する際、その経路上で障害が発生し、データ転送ができなくなったとき、その経路上の障害の回復を待ってデータ転送を再開させるか、SAN及びWAN上の他の経路を経由して、バックアップを完了させることとなる。また、IPを用いてデータ転送を行う従来技術（特許文献2）では、IPを用いたネットワーク内で、障害を回避するためのネットワーク経路変更を早く実行する。SANやIPを用いてデータ転送を行う従来技術（特許文献3）では、ネットワークの使用環境に応じてネットワーク指定を可能とする。

【0013】

しかし、これらの従来技術では、転送プロトコルの異なるネットワークを用いたデータ転送を想定していない、若しくは、形式的には転送プロトコルの異なるネットワークを用いたデータ転送は想定しているが、転送されるデータの転送単位については十分な考慮がなされていない。実際に、ストレージシステム間のデータ転送では、データの転送単位も考慮する必要がある、ストレージシステム間のデータを、ファイバーチャネル等を使用してSAN経由で転送する場合、SCSI (Small Computer System Interface) 等に基づいてブロック単位でデータを転送することとなる。しかし、サーバがIPを使用してLAN経由でデータ転送する場合、サーバが認識しているファイルシステムのファイル単位となる。従って、SAN及びWANの経路で、ストレージシステム間のデータ転送に失敗したデータ（ブロック単位）は、そのままでは、LAN及びWAN経由のサーバ間でデータ転送できない。この課題は、転送プロトコルを、単にSCSIからTCP/IP (Transmission Control Protocol / Internet Protocol) に変換しただけでは解決されず、データの転送単位についてどのように取扱えば良いのかを考慮する必要がある。特に、ブロック単位でのデータ転送処理の途中で障害が発生した場合の取扱いは、システム運用の際には重要である。

【課題を解決するための手段】

【0014】

本発明のコンセプトの一つは、ある記憶部（例えば、ストレージ、ディスク、ボリューム、記憶領域等）に格納されたデータを他の記憶部へ転送する際、ブロック単位のデータ転送とファイル単位のデータ転送とを適宜切り換えて利用することである。また、記憶部に格納されたデータをブロックレベルで管理するシステムの転送技術とファイルレベルで管理するシステムの転送技術とを適宜利用することである。これらにより、単一プロトコルでのデータ転送に制限されない柔軟性のあるデータ転送を実現する。また、データ転送の際の様々な条件・状況（ユーザニーズ、ネットワーク使用状況等）に応じたネットワーク選択の幅を拡張することができる。

【0015】

より具体的には、第1の記憶部から第2の記憶部へネットワーク（例えば、SAN、LAN、WAN等）を介してデータ転送を行う計算機システムであって、前記第1の記憶部に格納されたデータをブロック転送プロトコル等を用いてブロック単位で前記第2の記憶

部へ転送する処理を行う第1の制御部と、複数のブロックのデータからなるファイルと当該ファイルを構成するデータのブロックとを対応付けて管理するテーブル（例えば、ファイル／ブロック管理テーブル）と、前記第1の制御部からブロックを特定する情報（例えば、論理的なブロックアドレス）を受け、前記テーブルを用いて当該ブロックに対応するファイルを特定し、当該特定されたファイルをファイル転送プロトコル等を用いてファイル単位で前記第2の記憶部へ転送する処理を行う第2の制御部とを備えたことを特徴とする。ここで、第1の制御部と第2の制御部は、SANで構成されるように、ストレージシステムとファイル管理機能を有するサーバとに存在しても、NAS（Network Attached Storage）の構成のように同一システム・筐体に存在しても良い。前記ブロック転送プロトコルを用いた転送はファイバチャネル等によって各リソースが接続されているSAN、前記ファイル転送プロトコルを用いた転送はLANといったように、ネットワーク障害に柔軟に対応したり、ネットワーク特性を有効利用（比較的高速・高信頼・高価なネットワークと比較的低速・低信頼・安価なネットワークとの併用等）すべく複数のネットワークを介して行うようにしても良い。

【0016】

第1の制御部は、第1の記憶部に格納されたデータをブロック転送プロトコルを用いて転送処理に失敗した場合、例えば、第1の記憶部からのデータ読出しの障害やネットワーク障害を含む転送処理の障害を検知（障害と判断）した場合や、第2の記憶部を管理するシステムからデータ受信拒絶を受けた場合には、転送障害等に関するブロックを特定する情報を前記第2の制御部へ通知するようにしても良い。これにより、ユーザ設定等により予め想定したタイミングでのデータ転送処理の切換えの他、特にデータ転送処理に失敗した場合のデータ転送処理を担保可能となる。

【0017】

特定されたファイルが、前記転送障害に関するブロック以外のブロックのデータ（第1の制御部によるブロック転送プロトコルを用いた転送処理が完了したデータや転送処理が未実行のデータ）を含むことにより、ブロック単位の転送からファイル単位の転送への切換えを確実に行うことができる。

【0018】

第2の制御部は、前記ファイル単位での転送処理に失敗した場合、例えば、ファイルに対応するデータ読出しの障害やネットワーク障害を含む転送処理の障害を検知（障害と判断）した場合や、第2の記憶部を管理するシステムからデータ受信拒絶を受けた場合には、当該転送処理に失敗したファイルに関連する複数のブロックを特定し、当該複数のブロックのデータ転送を第1の制御部へ指示するようにしても良い。これによりファイル単位の転送処理からブロック単位の転送処理への切換えを可能とする。

【0019】

ここで、第1の記憶部は、少なくとも正ボリュームと副ボリュームとからなり、第1の制御部は、副ボリュームに格納されたデータのブロック単位での転送処理に失敗した場合、転送処理に失敗したデータのブロックを特定する情報を前記第2の制御部へ通知し、第2の制御部から前記転送処理に失敗したファイルに関連する複数のブロックのデータ転送指示を受けて、正ボリュームに格納された前記複数のブロックに対応するデータのブロック単位での転送処理を行うように運用することにより、通常正ボリュームを用いて実行される業務処理に極力影響を及ぼさないようなデータ転送手法も取り入れることが可能である。

【0020】

尚、上記データ転送手法は、バックアップ処理の際のデータ転送にも適用可能であり、本発明は、上述した機能を実現するプログラム若しくはプログラムを格納した記録媒体にも適用可能である。

【発明の効果】

【0021】

本発明によれば、複数のネットワークプロトコルを用いた柔軟性のあるデータ転送を、

データ転送単位が考慮された形で実現可能である。

【発明を実施するための最良の形態】

【0022】

以下本発明の実施の形態を、図面を参照しながら詳細に説明する。図1は、本発明が適用されるネットワーク環境の一実施形態を示す全体構成図である。

【0023】

本実施形態は、本発明のデータ転送手法を、データバックアップ処理を行うシステムに適用したものである。現用系のストレージシステム110と現用系のサーバ装置107は、ファイバーチャネルインターフェース（FC I/F）103を介してSAN106に接続されている。サーバ装置107は、システム全体の制御及びプログラム処理を行うマイクロチップなどの制御部、プログラムやデータを保持する記憶部を含んで構成されている。記憶部はメモリ等の主記憶装置やハードディスク等の補助記憶装置、データベース等から構成され、記憶部にはOS（Operating System）、ファイル転送プログラム、ファイル管理プログラム、業務プログラムの他、サーバ装置107の処理上保持すべきデータやプログラムが記憶されている。記憶部に格納されたプログラム等と制御部とが連動して、下述する各機能が実現される。

【0024】

サーバ装置107は、制御部上でファイル転送プログラムやOSが動作することにより、ファイル単位の転送処理を実行するファイル転送部108と、制御部上でファイル管理プログラムやOSが動作することにより、ファイル名とファイルを構成するデータとの対応付けを、後述するファイル/ブロック管理テーブル（サーバ装置107の内部に存在しても、外部であっても良い）等を用いて管理するファイル管理部109とを制御部の一機能として提供する。

【0025】

サーバ装置では、通常、UNIX（登録商標）等のOSと業務プログラムが実行されており、業務プログラムが使用するデータ（ファイル）は、ファイル/ブロック管理テーブル等を用いてブロックアドレスを特定してストレージシステム110へ通知され、ストレージシステム110の制御部であるディスクコントローラ111を介して、ボリューム112に格納される。ボリューム112は、単数の場合、複数の場合が考えられ、例えば、同じ内容のデータを複数のボリューム（正ボリューム、副ボリューム）に格納し、業務プログラムでは正ボリュームに格納されたデータを使用し、バックアップやリストアの際には副ボリュームに格納されたデータを使用するようにしても良く、サーバ装置107上のOSが採用しているファイルシステム形式でフォーマットされている。

【0026】

サーバ装置107は、ネットワークインターフェース（Net I/F）102を介して、SAN106とは別のネットワークであるLAN101に接続されている。LAN101、SAN106は、それぞれWAN-A104、WAN-B105に接続されている。本例では、WAN-AとWAN-Bは別のネットワークとしているが、同一のWANの例にも適用可能であり、現用系とバックアップ系との間にWANは介在せず、現用系とバックアップ系が、同一のSANに接続されている例にも適用可能である。WAN-A104、WAN-B105には、バックアップ系のストレージシステム130とバックアップ系のサーバ127が、それぞれLAN121及びSAN126を経由して接続されている。バックアップ系は、現用系と同様に、ストレージシステム130とサーバ装置127が、ファイバーチャネルインターフェース（FC I/F）123を介して、SAN126で接続されている。

【0027】

本実施形態では、サーバ装置の制御部とストレージシステムの制御部とが別個独立のシステムに存在する構成で説明するが、NASのようにサーバ装置の制御部とストレージシステムの制御部とが同一のシステムに存在しても良く、ボリューム112等の記憶部だけ別個独立の構成であっても良い。

【0028】

また、データバックアップ処理を行うシステムを例としているため、現用系とバックアップ系の構成をとっているが、お互いが現用系の構成等、システム間のデータ転送に広く適用可能である。

【0029】

図2は、現用系のストレージシステムのデータをバックアップするときのデータの流れを示したものである。ストレージシステム210のディスクコントローラ211は、ボリューム212に格納されているデータを読み出し、ファイバーチャネルインターフェース203を介してSAN206に送出する。データは、WAN-B205を介して、バックアップ系のSAN226に送信され、ファイバーチャネルインターフェース223を介して、ディスクコントローラ231が受け取り、ボリューム232に格納する。このとき、データはブロック単位233で転送される。バックアップ系のディスクコントローラ231は、ボリューム232への転送データの格納が成功したとき、その旨を現用系のディスクコントローラ211に応答する。現用系のディスクコントローラ211は、この応答を受けたら、当該データのバックアップが成功したと判断する。ディスクコントローラ211は、データ送出後、一定時間以内にこの応答を受けない場合、SAN206、226及びWAN-Bを介した経路上に何らかの障害が発生し、データ転送に失敗したと判断する。ディスクコントローラ211は、データ転送に成功したか否かの成否情報を、当該データに関連する情報であるブロック番号と共に、サーバ207上のファイル管理部209に通知する。ファイル管理部209は、この成否情報をファイル／ブロック管理テーブル（後述）で管理する。

【0030】

尚、ディスクコントローラ211は、データ転送に失敗した場合のみ、当該転送に失敗したデータを特定する情報（ブロック番号等）を、サーバ207上のファイル管理部209に通知しても良い。

【0031】

また、データ転送処理の失敗に、上述した転送経路上の障害の他、ディスクコントローラ211がボリューム212からデータを読み出す際の障害や、ディスクコントローラ231からデータ受信の拒絶を受けた場合を含めて運用することも可能であり、データ読み出し障害やデータ受信の拒絶の場合を、上述した転送経路上の障害と分けて運用しても良い。

【0032】

図3は、図2に示したWAN-B経由でのデータ転送で、障害により転送処理に失敗したデータブロックが発生した時、当該障害を回避してデータ転送及びバックアップを完了させるデータ転送の流れを示したものである。このデータ転送は、サーバ装置307上のファイル転送部308が行う。ファイル管理部309がファイル／ブロック管理テーブル（後述）をサーチし、転送成否フラグが失敗（Failure）になっているブロックを含むファイルを検知し、当該ファイル名をファイル転送部308に通知することで、ファイル転送部308は、転送が完了していないファイルを検知する。そのファイルをストレージシステム310のディスクコントローラ311を介してボリューム312から読み出し、ネットワークインターフェース302を介して、ネットワークLAN301に送出する。このときのデータ転送はファイル単位333となる。転送データは、WAN-A、LAN321経由でバックアップ系のサーバ装置327のファイル転送部328が受信し、ファイル管理部329が、ファイバーチャネルインタフェース（FC I/F）323、SAN326、ディスクコントローラ331を経由して、バックアップ系のストレージシステム上のボリューム332に格納する。これにより、WAN-B経由でのデータ転送に失敗したデータのデータ転送及びバックアップが完了する。ファイル管理部309が、ファイル／ブロック管理テーブル401をサーチするタイミングは、ストレージシステム310のディスクコントローラ311からデータ転送に失敗した旨の通知があったときでもよいし、一定期間ごとに定期的であってもユーザ指示に応じて行ってもよい。

【0033】

図4 (a) (b) は、ブロック単位での転送を行う際のデータ構造とファイル単位での転送を行う際のデータ構造とを示す一例である。

【0034】

図4 (a) は、S A N等を介してブロック単位のデータ転送を行う際の基本単位であるフレームの一例である。フレーム400は、フレームの境界を示すスタートオブフレームS O F (Start Of Frame) 401、エンドオブフレーム (End Of Frame) 405と、フレーム受信側のアドレス識別子やフレーム送信側のアドレス識別子を含むフレームヘッダ402、実際に転送される目的となるデータ部分であるデータフィールド403、サイクリックリダンダンシチェックC R C (Cyclic Redundancy Check) 404で構成される。データフィールド403は、0～2112バイトの間で可変であり、ブロック単位のデータ転送を実現する。図4 (b) は、L A N等を介してファイル単位のデータ転送を行う際のフレームの一例である。フレーム410は、フレームヘッダ411の他、送信元アドレスや送付先アドレスを特定するI Pアドレスを含むI Pヘッダ412と、送信データの送り側アプリケーションのストリーム中でのバイト位置を示すシーケンス番号や相手からどこまでのデータを受取ったかをバイト単位で示す確認応答番号といったデータ転送の信頼性を確保するための情報を含むT C Pヘッダ413と、実際に転送される目的となるデータ部分であるデータフィールド414を含んで構成され、ファイル単位のデータ転送を実現する。T C P / I Pで通信相手先やアプリケーションを特定し、F T P (File Transfer Protocol) 等を用いてファイル転送開始フラグやファイル転送終了フラグを合図にファイル単位での転送を行うことにより、ファイル単位のデータを一括して転送若しくは固定データ長で分割して順次転送する。分割して転送する場合は、受信側で転送されるデータがファイル単位のデータとして再構成可能なように、シーケンス番号等で管理される。

【0035】

図5は、ファイル／ブロック管理テーブルを示す。ファイル／ブロック管理テーブル501では、ファイル名502等のファイル識別子毎に、当該ファイルの内容を格納しているボリューム上のブロックのブロック番号503等といったブロックアドレスを特定する情報を関連付けて管理している。ブロック番号は、ブロック毎に付された番号であり、ボリューム上の位置情報を含むものとする。ストレージシステム110、130において物理ボリュームの物理的なブロックアドレスに対応させて、論理ボリュームの論理ブロックアドレスを管理している場合 (例えば、ディスクコントローラ111、131が物理ブロックアドレスと論理ブロックアドレスとの対応付けを管理している場合) は、論理ブロックアドレスをブロック番号としても良い。順番504は、ファイルの内容が書き込まれているブロックの順番である。また、ブロック番号503毎に転送成否フラグ505を関連付けている。図5の場合、ファイル“/tmp/work.dat”の内容は、ブロック番号001から003までのブロックに、001, 002, 003の順で格納されており、当該全ブロックは、転送が成功したこと (Success) が示されている。一方、ブロック番号006のブロックは、転送が失敗 (Failure) したことを示している。つまり、ファイル“/tmp/library.dat”の内容の一部が転送されていないことを示している。また、ファイル“/tmp/sample.txt”は、転送が未実施 (Undone) であることを示している。ファイル管理部209は、ディスクコントローラ211から送信された転送成否情報に応じて、ファイル／ブロック管理テーブル501の転送成否フラグ505を更新する。

【0036】

尚、先にも述べたとおり、データ転送に失敗した際のみ、当該転送に失敗したデータを特定する情報 (ブロック番号等) を、サーバ装置上のファイル管理部に通知する場合等では、転送処理が失敗した場合のみフラグを立てて管理しても良い。データ転送処理の失敗に、上述した転送経路上の障害の他、ディスクコントローラがボリュームからデータを読み出す際の障害や、ディスクコントローラからデータ受信の拒絶を受けた場合を含めて運用する場合は、同様に転送処理が失敗したものとして取扱えばよく、データ読出し障害やデータ受信の拒絶の場合を、上述した転送経路上の障害と分けて運用する場合は、失敗内

容が特定できる情報をファイル／ブロック管理テーブルに追加して管理しても良い。

【0037】

図6は現用系のストレージシステムのディスクコントローラが実行する処理のフローチャートである。サーバ装置等から命令を受信すると、ディスクコントローラではこの命令の内容を判定する(ステップ602)。命令が、ファイルシステム構築命令の場合、ステップ603でボリューム上にファイルシステムを構築する(新たな領域確保、ファイルシステムでの初期化等)。このタイミングと関連して、サーバ側では、ファイル／ブロック管理テーブルが生成される(図7のステップ704参照)。命令が、ファイルの生成・更新・削除の場合、ステップ604で、命令に応じて、ボリューム上にブロック単位でデータを書き込んだり、削除したりする。ステップ605では、ファイルの生成・更新・削除に応じたボリューム上のブロック番号をサーバ装置上のファイル管理部に通知する。これにより、通常の業務運用中でのファイル操作の結果を、サーバ上のファイル／ブロック管理テーブルに反映することができるようになる(図7のステップ706参照)。

【0038】

命令がバックアップの場合、ボリュームからデータをブロック単位で読み込む(ステップ606)。ステップ607で、読み出したデータをブロック単位で、ファイバチャネル・インタフェース(FC I/F)に送出する。ステップ608では、バックアップ系ストレージシステムからのデータ受信応答を待つ。受信応答があった場合、データ転送が成功したと判断し、ステップ609で、その旨をサーバ装置上のファイル管理部に、当該データのブロック番号と共に通知する。一定時間経っても通知がない場合、ステップ610で、サーバ装置上のファイル管理部に転送失敗したことをブロック番号と共に通知する。これにより、バックアップの結果をサーバ装置上のファイル／ブロック管理テーブルに反映することができるようになる(図7のステップ708参照)。ステップ611で、転送及びバックアップすべきデータがボリュームに残っているかどうかを判定する。残っている場合、ステップ606に戻り、次のデータをボリュームから読み出す。残っていない場合、転送及びバックアップ処理は終了する。

【0039】

尚、ディスクコントローラは、データ転送に失敗した場合のみ、当該転送に失敗したデータを特定する情報(ブロック番号等)を、サーバ装置上のファイル管理部に通知する場合は、ステップ609の処理を省略可能であり、ステップ610ではブロック番号の通知をもって、転送処理に失敗したとサーバ装置側で判断させるようにしても良い。また、データ転送処理の失敗に、ディスクコントローラがボリュームからデータを読み出す際の障害を含めて運用する場合は、ステップ606の処理の失敗をサーバ装置上のファイル管理部に通知すれば良く、バックアップ系ストレージシステムからデータ受信の拒絶を受けた場合を含めて運用する場合は、ステップ607の処理に対するその旨の応答があったことをサーバ装置上のファイル管理部に通知すれば良い。データ読み出し障害やデータ受信の拒絶の場合を、上述した転送経路上の障害と分けて運用する場合は、失敗内容が特定できる情報を含めて上記通知を各々行えば良い。

【0040】

図7は、現用系のファイル管理部が実行する処理のフローチャートである。サーバ装置と接続されたクライアント装置、サーバ装置内の業務プログラム等から命令を受信すると、この命令の内容を判定する(ステップ702)。命令がファイルシステム構築の場合、ストレージシステム上のディスクコントローラにファイルシステム構築命令を出し(ステップ703)、ファイル／ブロック管理テーブルを生成する(ステップ704)。命令が、ファイル生成・更新・削除の場合、ディスクコントローラに対して、当該命令に応じたデータ操作の命令を出し(ステップ705)、ファイル生成・更新・削除に応じて、ファイル／ブロック管理テーブルを更新する(ステップ706)。

【0041】

命令がバックアップの場合、ディスクコントローラにバックアップ命令を出し、ディスクコントローラから、転送成否情報やブロック情報を受信する(ステップ707)。ステ

ップ708で、転送成否情報やブロック情報に応じて、ファイル／ブロック管理テーブルを更新する。サーバレスバックアップ方式や、バックアップサーバを別個に構築する場合は、ディスクコントローラにバックアップ命令を直接出さず、ディスクコントローラから転送成否情報やブロック情報を受信した場合にファイル／ブロック管理テーブルを更新する処理としても良い。

【0042】

命令がファイル転送の場合、ステップ709で、ファイル／ブロック管理テーブルをサーチし、転送に失敗したブロックに関連のあるファイルを検出する。関連があるファイルとは、そのファイルの内容の一部又は全部が当該ブロックに格納されているファイルのことである。図5の例の場合、ブロック番号006が転送に失敗しており、それに対応するファイル“/tmp/library.dat”が検出されるファイルとなる。ここで、転送経路上の障害と、データ読出し障害やデータ受信の拒絶の場合を分けて運用する場合は、失敗内容が特定できる情報がファイル／ブロック管理テーブルに追加されているので、転送経路上の障害によるブロック転送失敗データに関するファイルのみをファイル転送する、データ受信拒絶によるブロック転送失敗データに関するファイルのみをファイル転送の対象としないといったファイル転送ポリシーを予め設定してファイル管理部で管理する、若しくはファイル転送命令の際に当該転送ポリシーを追加する等により、データ転送失敗の内容に応じた柔軟な運用も可能となる。

【0043】

ステップ710で、検出したファイル名と当該ファイルに関するファイル／ブロック管理テーブルの情報（ブロック番号・順番）をファイル転送部に通知する。ステップ711で、バックアップ系サーバから受信したファイル書込み成否情報に応じて、ファイル／ブロック管理テーブルを更新する。例えば、“/tmp/library.dat”の書込みが成功した旨の情報を受取ったときは、図5に示すファイル／ブロック管理テーブルのブロック番号006の転送成否フラグをSuccessに変更する。

【0044】

ファイル転送処理のタイミングが、ストレージシステムのディスクコントローラからデータ転送処理に失敗した旨の通知があったときとする場合は、クライアント装置、サーバ装置内の業務プログラムからの指示を受けることなく当該通知を契機として、ステップ709で、ディスクコントローラから通知される情報に基づいて、転送処理に失敗したブロックに対応するファイルを特定（未転送ファイルを検出）する処理を行い、ステップ710へ移行しても良い。

【0045】

また、ステップ711でファイル書込みに失敗した旨の情報を受取ったとき等、ファイル転送処理に失敗した場合、当該転送処理に失敗したファイルに対応するブロックをファイル／ブロック管理テーブルを用いて特定し、当該ブロックのデータ転送をストレージシステムのディスクコントローラに指示しても良い。これにより、ネットワーク状況に応じたデータ転送手法の選択が可能となる。さらに、ファイル転送処理の失敗が、ディスクコントローラからデータ転送処理に失敗した旨の通知があった後のファイル転送処理における場合は、ディスクコントローラと連携して、ディスクコントローラがデータ転送処理に失敗した旨の通知をファイル管理部に出した際の当該データ転送処理に利用したボリュームと異なるボリュームを用いてデータ転送処理を行うようにしても良い。例えば、ディスクコントローラとファイル管理部との間でやりとりされるブロック番号が論理ブロックアドレスであり、ディスクコントローラで当該論理ブロックアドレスに対応付けて正ボリュームと副ボリュームの物理ブロックアドレスを管理している場合、最初のブロック単位のデータ転送処理は、副ボリュームから行い、データ転送処理に失敗した場合はファイル単位のデータ転送処理を行い、さらにデータ転送処理に失敗した場合は正ボリュームを用いたブロック単位のデータ転送処理を行うことによって、ネットワーク状況に応じたデータ転送手法の選択をしながら、正ボリュームを利用した業務処理等があった場合に、当該処理を極力妨げないような構成とすることができる。

【0046】

図8は、現用系のファイル転送部が実行する処理のフローチャートである。ファイル転送部におけるファイル転送処理では、ステップ803で、ファイル管理部から転送すべきファイル情報を得る。このファイル情報は、図7のステップ709、ステップ710によって得られる情報等である。当該ファイル情報（未転送ファイル名、ファイル／ブロック管理テーブルの情報等）に基いて、転送すべきファイルを構成するブロック番号を指定して問い合わせることによりストレージシステムのディスクコントローラからブロック番号に対応するデータを得て、ファイル単位のデータとして構築する。例えば、ファイルを構成するデータにファイル名を付す。

【0047】

尚、ファイル管理部からファイル転送部へ通知されるファイル情報にファイルを構成するデータを含めても良い。その場合、転送すべきファイルを構成するデータを、ブロック番号を指定してストレージシステムのディスクコントローラへ問い合わせるのはファイル管理部となる。

【0048】

ステップ804で、転送すべきファイルと当該ファイルに関連するファイル／ブロック管理テーブルの情報（ブロック番号・順番）をネットワーク・インタフェース（Net I/F）・LANを経由してバックアップ系のサーバ装置に転送する。ファイル管理部から受けた転送すべきファイル情報から、ブロック単位でのデータ転送処理に失敗したブロックに関連のあるファイル単位のデータを構築し、当該ファイル単位でデータを転送することになるので、すでにブロック単位でのデータ転送処理に成功しているデータも含めて、ファイルを構成するデータの先頭から転送されることになる。ファイル単位のデータ転送処理によって、すでにブロック単位のデータ転送処理に成功しているデータについてはバックアップ系のボリュームで上書きされることとなる。

【0049】

ステップ805で、バックアップ系のサーバ装置からファイル書込みの成否情報を受信する。これは、図12のステップ1205に対応したものである。ステップ806では、ファイル書込みの成否情報をファイル管理部に通知する。これにより、図7のステップ711で、ファイル管理部は、ファイル転送の成否情報に応じて、ファイル／ブロック管理テーブルを更新できる。

【0050】

図9は、バックアップ系のディスクコントローラが実行する処理のフローチャートである。ディスクコントローラ部におけるボリューム書込み処理では、ステップ903で、当該書込みがファイル単位の書込みであるか、ブロック単位の書込みであるかを判定する。つまり、ストレージシステム間のブロック単位のデータ転送による書込みであるか、サーバ装置を経由したファイル転送による書込みであるかを判定する。ブロック単位の書込みである場合、ステップ904でボリュームにデータを書込み、ステップ905で書込み成否情報を現用系のストレージシステム上のディスクコントローラに通知する。ファイル単位の書込みである場合、ステップ906でファイルのデータをブロック単位に分割する。例えば、ファイル単位のデータ内にブロックの区切りの情報が付加されている場合は、このブロックの区切りの情報を利用して分割したり、ファイル単位のデータの先頭から1ブロックのサイズ毎に分割したりする。ステップ907で分割されたデータを現用系から転送されたファイル／ブロック管理テーブルのブロック番号と順番に合わせて、ボリュームに書込む。ステップ906とステップ907を具体例をあげてより詳細に説明する。転送されたファイルが、図5に示す“/tmp/library.dat”であるとする。このファイルは、現用系では、ブロック番号004、006、005の3つブロックに、この順番で格納されている。ステップ906でこのファイルが分割されると、3つのデータに分割される。ステップ907で、この3つのデータをボリューム上のブロック番号004、005、006の位置に、004、006、005の順番に書込む。ステップ908で、書込み成否情報をサーバ装置上のファイル管理部に通知する。

【0051】

尚、図9のバックアップ系のディスクコントローラにおけるボリューム書込み処理は、ブロック単位の書込み処理のみとしても良い。本実施例を図10を用いて説明する。本実施例の場合、ファイル単位のデータをブロック単位へ分割する処理は、バックアップ系のサーバ装置における制御部で行うことになる。従って、図9のステップ903に相当する判定処理は省略可能であり、ステップ1001で、ストレージシステム間のブロック単位のデータ転送による書込み、またはサーバ装置からのブロック番号と当該ブロック番号に対応するデータの指定を受けた書込みを実行する。書き込み成否情報は、ステップ1002で、ボリュームへの書き込みが、現用系のディスクコントローラからのものか、バックアップ系のサーバ装置からのものかを判断し、前者については現用系のディスクコントローラへ通知し（ステップ1003）、後者についてはサーバ装置のファイル管理部へ通知する（ステップ1004）こととなる。

【0052】

図11は、バックアップ系のファイル管理部が実行するステップのフローチャートである。ファイル管理部におけるファイル受信処理では、ステップ1103でファイル転送部が受信したファイルをボリュームに書き込む命令をディスクコントローラに出す。ディスクコントローラがファイル単位のデータをブロック単位へ分割する処理を行う場合、図9のステップ903、906、907を介してボリュームにデータが書き込まれる。そして、ステップ1104で、ディスクコントローラから書き込み成否情報を受取る。これは、図9のステップ908に対応したものである。サーバ装置においてファイル単位のデータからブロック単位のデータに変換する場合は、ステップ1103の処理において、ファイルのデータをブロック単位に分割し、分割されたデータを現用系から転送されたファイル／ブロック管理テーブルのブロック番号と順番に合わせ、ブロック番号と当該ブロック番号に対応するデータの指定をしてディスクコントローラに書込み命令を出す。この処理は、ファイルを構成するデータ分繰り返す。ディスクコントローラから当該データの書き込み成否情報を受取る（ステップ1104）。これは図10のステップ1004に対応したものである。ステップ1105で、当該書き込み成否情報をファイル転送部に通知する。

【0053】

図12は、バックアップ系のファイル転送部が実行するステップのフローチャートである。ファイル転送部におけるファイル受信処理では、ステップ1203で、LAN及びネットワークインタフェース（Net I/F）を経由して、現用系サーバからファイルと当該ファイルに関するファイル／ブロック管理テーブルの情報を受信する。ステップ1204で、ファイル管理部にファイルを受信した旨を通知する。これにより図11のステップ1103のファイル書込み命令がディスクコントローラに出される。ステップ1205で、書き込み成否情報をネットワークインタフェース（Net I/F）・LAN経由で現用系のファイル転送部に通知する。図8のステップ805はこれに対応したものである。

【0054】

このように、本発明の実施形態によれば、ストレージシステムに格納されたデータをバックアップする際に、SANを使用してデータ転送中に障害が発生してデータ転送ができなかったとき、SANとは別の既設のネットワーク（LAN等）とサーバ装置を使用してデータ転送をすることができる。SANを使用したデータ転送を既設のネットワーク（LAN等）を使用したデータ転送で代替する場合、転送するデータの単位が異なることが問題となるが、サーバ上のファイル管理部がファイル／ブロック管理テーブルでファイルとブロックの関係を管理することにより解決する。

【図面の簡単な説明】**【0055】**

【図1】 本発明が適用されるネットワーク環境の一実施形態を示す全体構成図

【図2】 現用系のストレージシステムのデータをバックアップするときのデータの流れの一例を示した図

【図3】データ転送処理の障害を回避してバックアップするときのデータの流れの一例を示した図

【図4】(a)ブロック単位での転送を行う際のデータ構造を示す一例、(b)ファイル単位での転送を行う際のデータ構造を示す一例

【図5】ファイル/ブロック管理テーブルの一例を示す図

【図6】現用系のディスクコントローラが実行するステップのフローチャートの一例を示す図

【図7】現用系のファイル管理部が実行するステップのフローチャートの一例を示す図

【図8】現用系のファイル転送部が実行するステップのフローチャートの一例を示す図

【図9】バックアップ系のディスクコントローラが実行するステップのフローチャートの一例を示す図

【図10】バックアップ系のディスクコントローラが実行するステップのフローチャートの他の一例を示す図

【図11】バックアップ系のファイル管理部が実行するステップのフローチャートの一例を示す図。

【図12】バックアップ系のファイル転送部が実行するステップのフローチャートの一例を示す図。

【符号の説明】

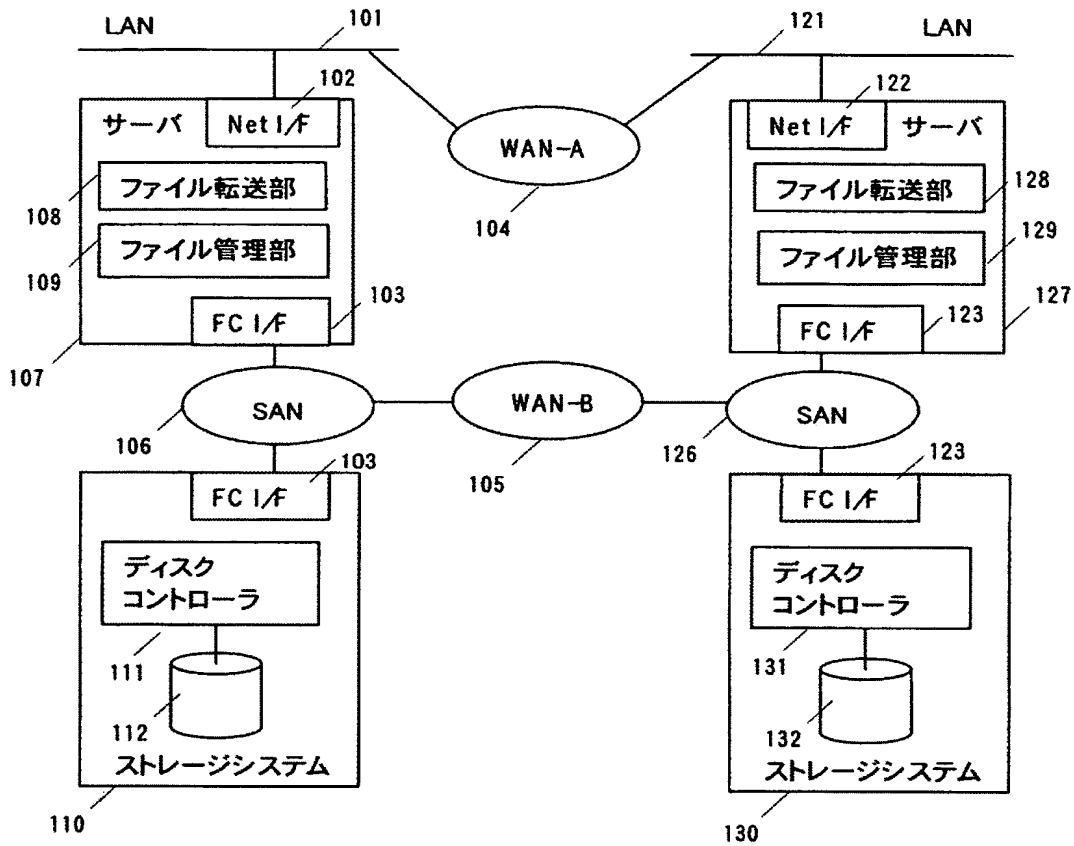
【0056】

101：現用系のLAN（ローカルエリアネットワーク）、102：ネットワークインターフェース、103：ファイバーチャネルインタフェース、104：現用系のLANとバックアップ系のLANを介在するWAN（ワイドエリアネットワーク）、105：現用系のSANとバックアップ系のSANを介在するWAN（ワイドエリアネットワーク）、106：現用系のSAN（ストレージエリアネットワーク）、107：現用系のサーバ装置、108：現用系のファイル転送部、109：現用系のファイル管理部、110：現用系のストレージシステム、111：現用系のディスクコントローラ、112：現用系のボリューム、121：バックアップ系のLAN（ローカルエリアネットワーク）、122：ネットワークインターフェース、123：ファイバーチャネルインタフェース、126：バックアップ系のSAN（ストレージエリアネットワーク）、127：バックアップ系のサーバ装置、128：バックアップ系のファイル転送部、129：バックアップ系のファイル管理部、130：バックアップ系のストレージシステム、131：バックアップ系のディスクコントローラ、132：バックアップ系のボリューム

【書類名】 図面

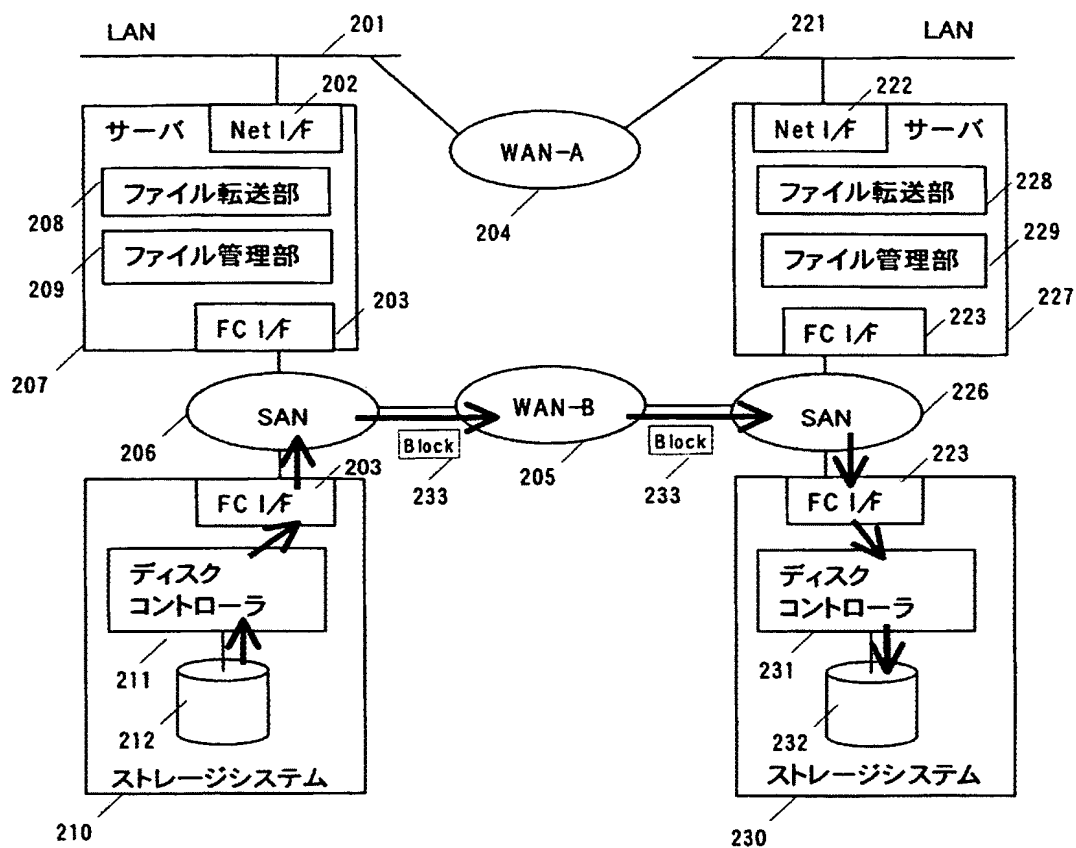
【図 1】

図 1



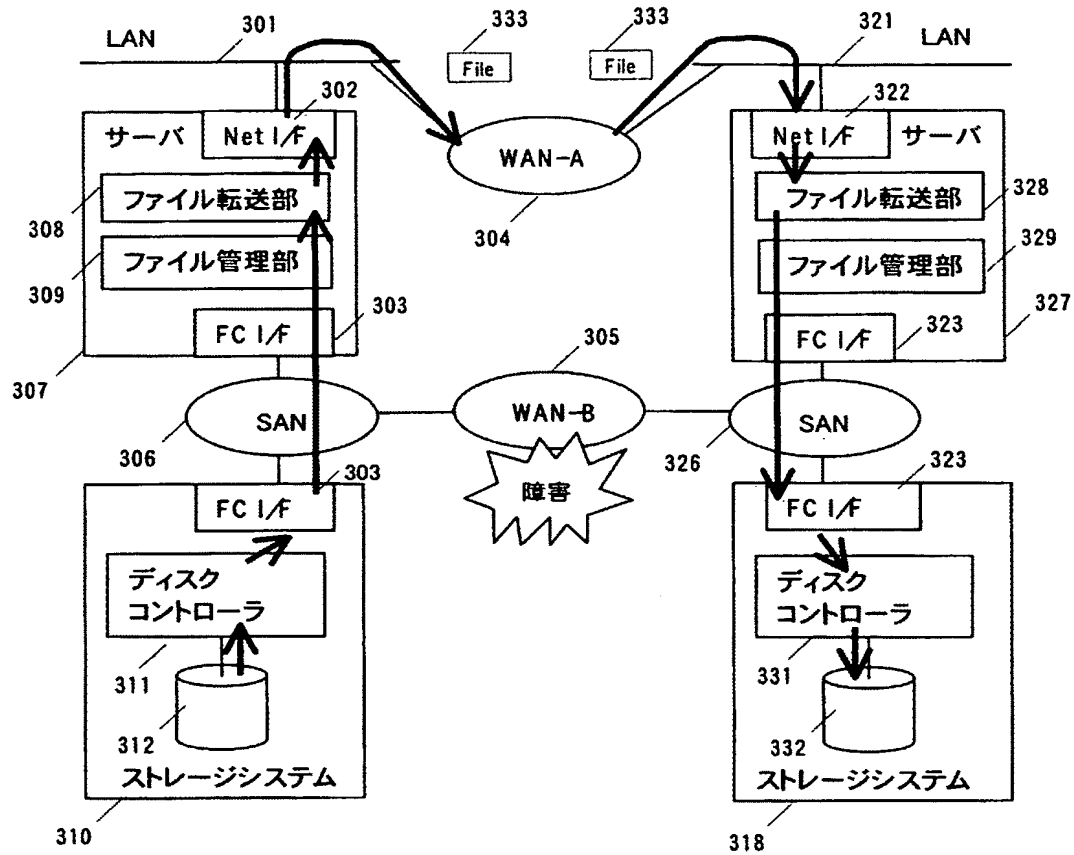
【図 2】

図 2



【図 3】

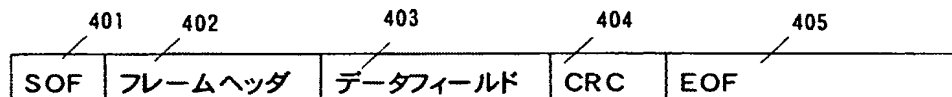
図 3



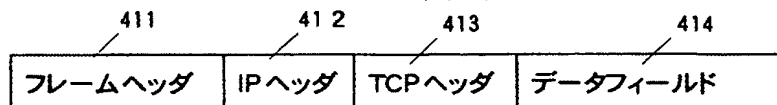
【図 4】

図 4

(a) ブロック転送



(b) ファイル転送



【図 5】

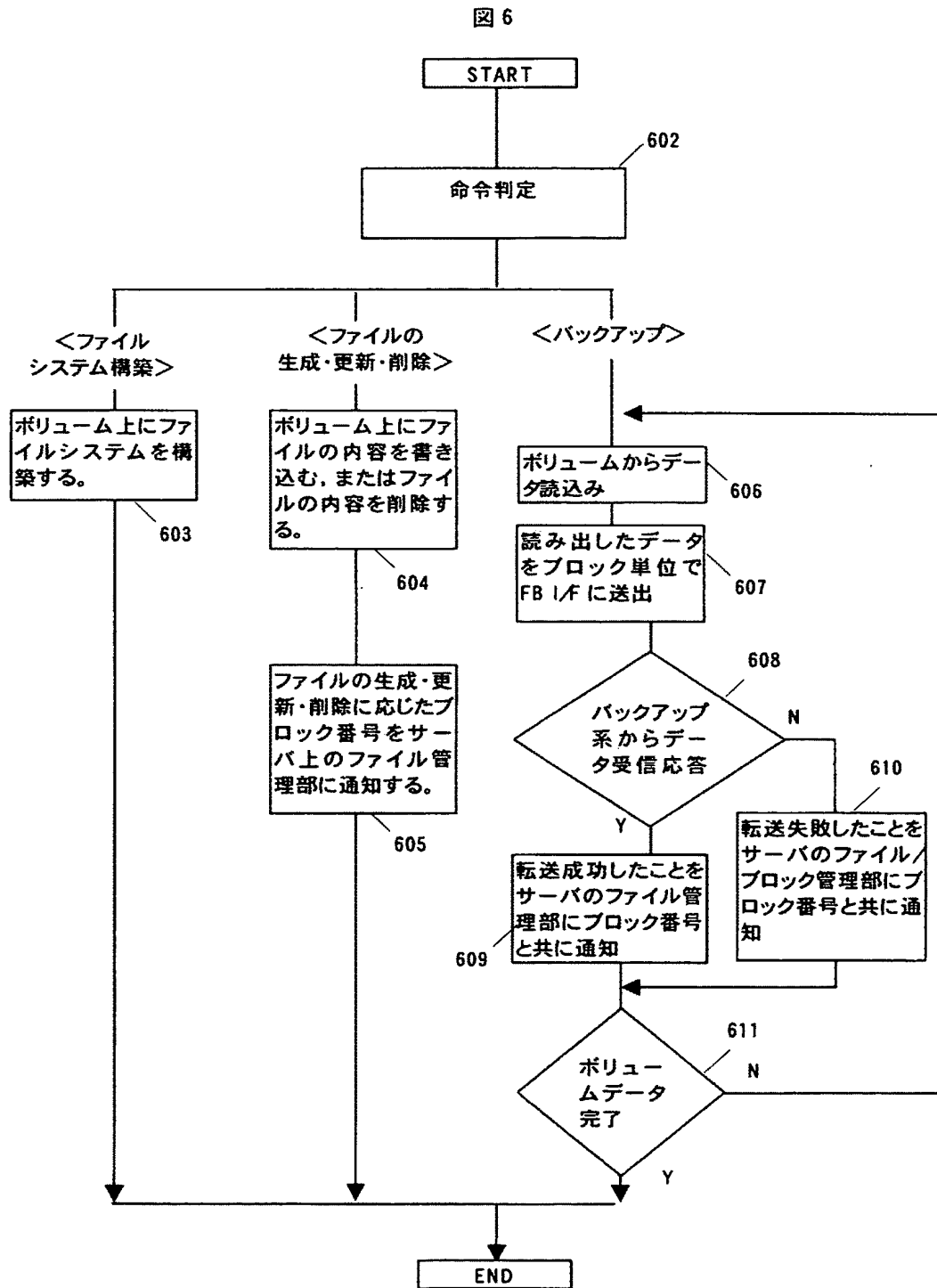
図 5

502 ファイル名	503 ブロック番号	504 順番	505 転送成否フラグ
/tmp/work.dat	001	1	Success
/tmp/work.dat	002	2	Success
/tmp/work.dat	003	3	Success
/tmp/library.dat	004	1	Success
/tmp/library.dat	005	3	Success
/tmp/library.dat	006	2	Failure
/tmp/sample.txt	007	1	Undone

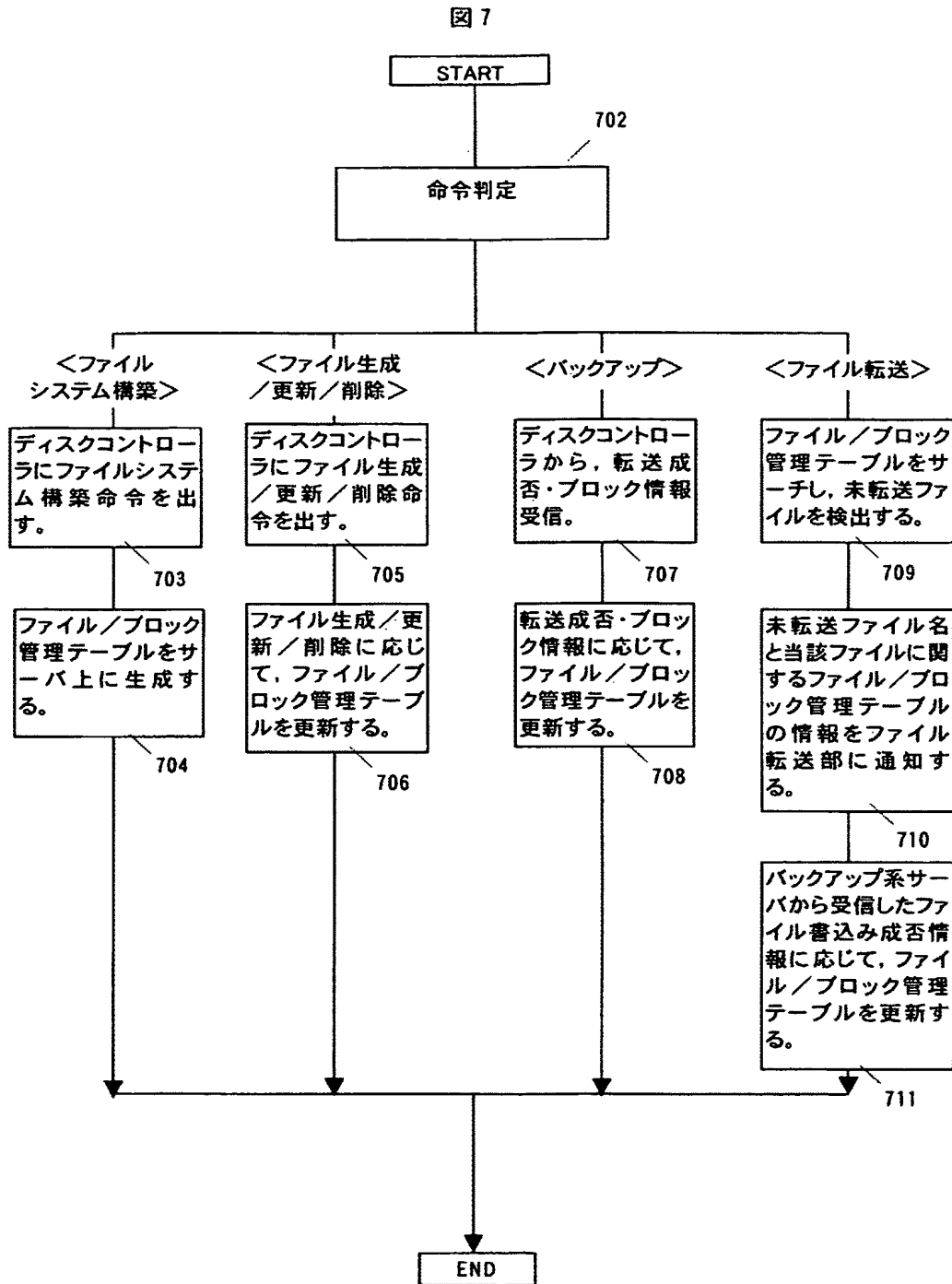
501

ファイル/ブロック管理テーブル

【図 6】

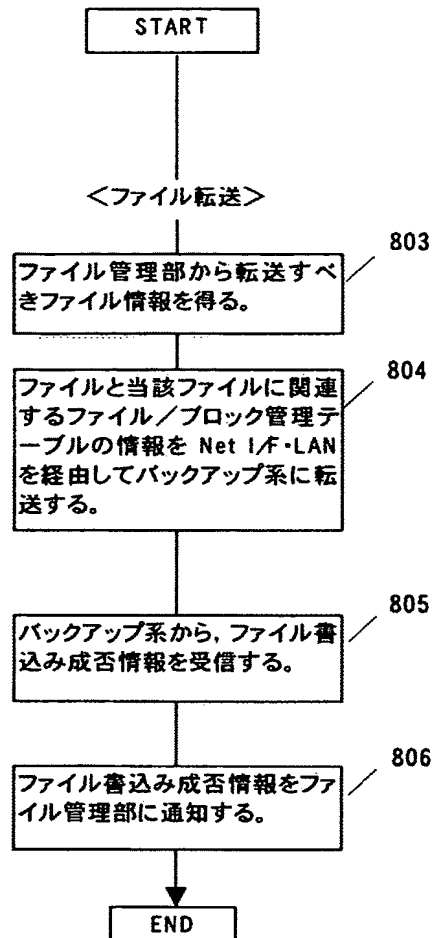


【図 7】



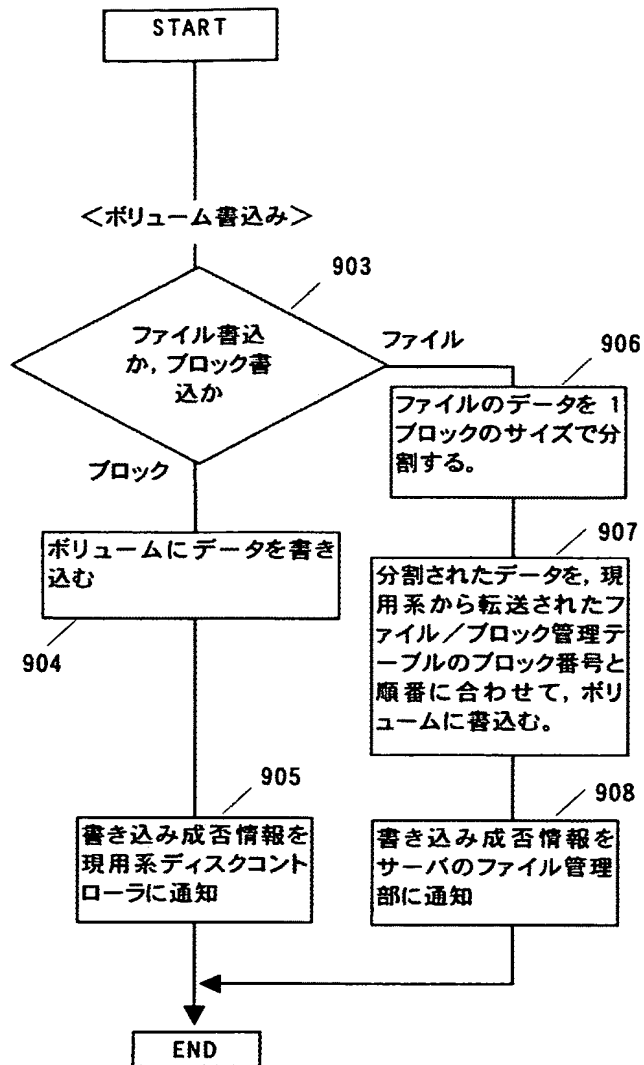
【図 8】

図 8



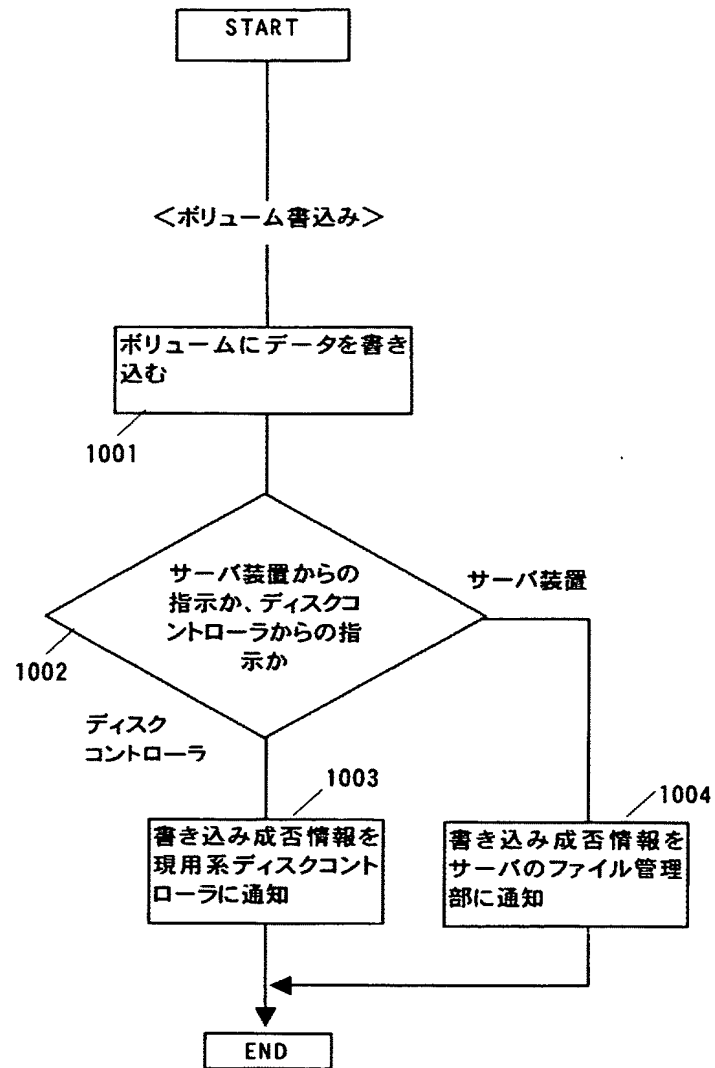
【図 9】

図 9



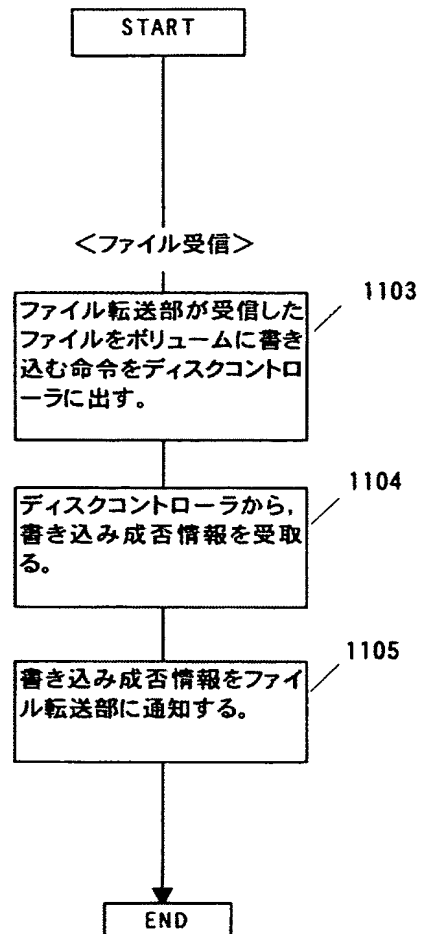
【図 10】

図 10



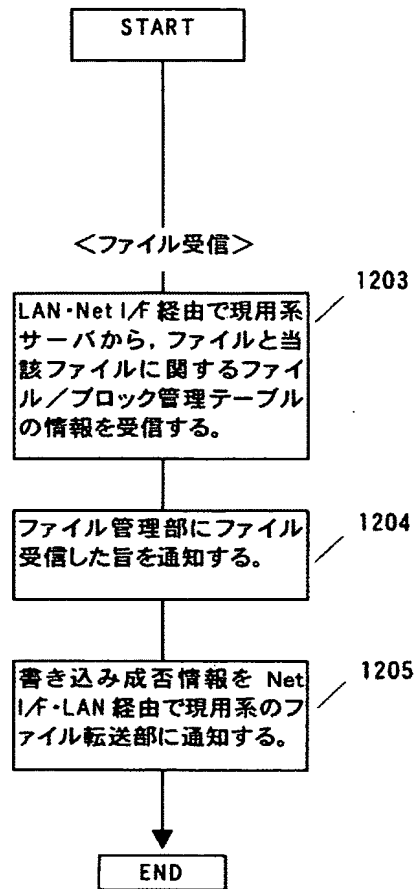
【図 11】

図 11



【図 12】

図 12



【書類名】 要約書**【要約】****【課題】**

計算機システムの記憶部に格納されたデータをネットワーク（S A N）を経由してデータ転送する際、障害等によりデータ転送が失敗したとき、別のネットワーク（L A N）を使用してデータ転送することで障害を回避する。

【解決手段】

第1の記憶部から第2の記憶部へネットワークを介してデータ転送を行う計算機システムであって、前記第1の記憶部に格納されたデータをブロック転送プロトコルを用いて前記第2の記憶部へ転送する処理を行う第1の制御部と、複数のブロックのデータからなるファイルと当該ファイルを構成するデータのブロックとを対応付けて管理するテーブルと、前記第1の制御部からブロックを特定する情報を受け、前記テーブルを用いて当該ブロックに対応するファイルを特定し、当該特定されたファイルをファイル転送プロトコルを用いて前記第2の記憶部へ転送する処理を行う第2の制御部とを備えたことを特徴とする計算機システム。

【選択図】 図 3

認定・付加情報

特許出願の番号	特願 2 0 0 3 - 4 2 6 1 6 8
受付番号	5 0 3 0 2 1 1 4 5 0 8
書類名	特許願
担当官	第七担当上席 0 0 9 6
作成日	平成 1 5 年 1 2 月 2 5 日

< 認定情報・付加情報 >

【提出日】 平成15年12月24日

特願 2 0 0 3 - 4 2 6 1 6 8

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 5 1 0 8]

1. 変更年月日	1 9 9 0 年 8 月 3 1 日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台 4 丁目 6 番地
氏 名	株式会社日立製作所